ESD-TR-65-461

(FINAL REPORT)

INTEGRATION OF CONCURRENT VISUAL AND AUDITORY MESSAGES

Donald B. Devoe

OCTOBER 1965

DECISION SCIENCES LABORATORY
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

ESRH

(Prepared under Contract No. AF 19 (628)-4073 by Sylvania Electronic Systems
A Division of Sylvania Electric Products Inc., 40 Sylvan Road, Waltham, Mass.)

DDC AVAILABILITY NOTICE

Copies have been deposited with the Defense Documentation Center. (DDC)

DISSEMINATION NOTICE

DDC release to the Clearing House for Federal Scientific & Technical Information is authorized. (CFSTI)

LEGAL NOTICE

OTHER NOTICES

Do not return this copy. Retain or destroy.

INTEGRATION OF CONCURRENT VISUAL AND AUDITORY MESSAGES

by

Donald B. Devoe

FOREWORD

Project 7682, Task 768201

This document is the final technical report under Contract AF19(628)-4073, between the Electronic Systems Division of the Air Force Systems Command and the Applied Research Laboratory of the Sylvania Electronic Systems Division, Sylvania Electric Products, Inc. The contract title is: "Investigation of Factors which Influence the Comprehensibility of Information Presented in Various Combinations of Visual and Auditory Form." Contract Monitor was Major William H. Watkins of ESD's Decision Sciences Laboratory (ESRHD). The research was conducted during the period 16 March 1964 to 15 July 1965 and this report is submitted on 15 August 1965.

The author wishes to acknowledge the invaluable assistance rendered by his two assistants: Miss Anne Flecchia and Mr. Barry Goldstein. Their aid in the selection of stimulus material, design of experiments, pretesting of procedures, collection and analysis of data, and interpretation of results was substantial and contributed significantly to the quality of the results. Our Contract Monitor and several of his colleagues on the staff of the Decision Sciences Laboratory were a regular source of help, and their contributions in both technical counsel and administrative assistance are gratefully noted. Finally, heartfelt thanks are extended to our experimental subjects, whose performance of the tasks imposed upon them constituted the real substance of the study.

This technical report has been reviewed and is approved.

Donald W. Connolly

DONALD W. CONNOLLY
Project Officer
Decision Sciences Laboratory

Roy Morgan

ROY MORGAN
Director
Decision Sciences Laboratory

ABSTRACT

Two experiments were performed involving the concurrent presentation to human subjects of two messages, one auditory and one visual, followed by a question requiring information from both messages. The results indicated that bimodally-presented information can be integrated for decision making. However, there was no evidence of an advantage to bimodal presentation as a means of unburdening an overloaded sense. The implications of the results for displays and communications in complex control centers are discussed and directions for future research are suggested.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

SECTION 1

INTRODUCTION

Although they may differ considerably in detail, the control centers of all large Air Force command and control systems are alike in containing both visual and auditory displays of information. Visual displays may vary from wall-sized situation boards, through individual cathode-ray-tubes or slide-projection devices, to posted teletype notices. Auditory displays may include radio and telephone communication systems, both individual (earphones) and group (loudspeakers), as well as various alerting and warning devices (buzzers, horns, and bells).

Do auditory and visual messages that are received together interfere with one another? Obviously, if serious and consistent interference occurred, bisensory inputs would have been abandoned early in the history of display development. We need only observe scope operators at work with their headsets on to realize that bisensory integration can and does occur. This does not mean that interference cannot also occur, however.

Assuming for the moment that integration of information received simultaneously through different senses does occur readily, is there any advantage to be gained from such presentation? First, if performance with data received through different senses is simply <u>no worse</u> than when only one sense is used, the interface between various communication systems and human operators is more flexible and can be designed to utilize the most convenient mode of presentation. That is, it will not be necessary to convert all auditory messages to visual or vice versa. Second, it has been suggested that, when one sense is overloaded with information, performance might be <u>better</u> if the information input is shared between two senses, thus <u>unburdening</u> one of them.

How do system designers provide for the most advantageous integration of visual and auditory information while minimizing interference? We would like to feel that there are certain guiding principles that effectively say: "This and this kind of information will integrate when presented through different senses, while this and this should be avoided". Unfortunately, the selection of what kind of information is presented to what sense with regard to sensory overload and chances of integration or interference is more a matter of tradition, convenience, or feasibility ... not because guiding principles are ignored, but because they are not available. So this area seems to warrant research.

The research reported here was an exploratory study aimed at determining whether or not two messages received together via different senses could be put together for decision making, whether bimodal presentation was advantageous, and what characteristics of the messages influenced this integrative capability. Although the desired guiding principles for determining when to employ bisensory stimulation and what to present to which sense were not expected to emerge in full clarity from our modest study, we did expect to verify or fail to verify some assumptions and to detect some principles in outline that might be worthy of further study.

SECTION II

BACKGROUND AND APPROACH

LITERATURE RELATED TO BISENSORY STIMULATION

From the earliest years of experimental psychology, there has been interest and experimentation in intersensory interactions. The early literature (which will not be reviewed here) showed considerable interaction among senses, particularly in terms of interference effects and both raising and lowering of thresholds in one sense by stimulation in another. Harris[1] reviewed the highlights of this literature, pointing out interesting behavioral and neurological parallels among the senses.

Following World War II, as the new field of Human Engineering focused attention on the need to consider human characteristics in the design of equipment for human use, concern arose as to the implications of intersensory effects for the design of communications systems and visual displays. A contract awarded to the University of Virginia by the U. S. Air Force for the study of vision and audition as communication channels organized the general problem area, applied the experimental literature to it, and produced several important summary reports.[2,3,4] Mowbray and Gebhard[5], in a report first appearing in 1956, revised in 1958, reprinted by Sinaiko[6], and liberally quoted in later guidebooks, extended the treatment in terms of information channels to all senses. London[7] reviewed the Russian literature. At about the same time, the San Diego State College Foundation[8] published a 245-item bibliography on sensory input channels.

Generally, both intersensory facilitation and inhibition were found in the studies reviewed, depending upon the various conditions being tested. It was possible to classify broadly the input

situations where each sensory channel could be used effectively. Within a single sense, considerable progress was made in determining how competing messages could be sorted out -- for example, Broadbent's work[9,10,11,12] on competing auditory messages. However, very little of the work reviewed was directly relevant to the integration of information meant to be put together but presented via different senses.

In the field of vigilance and warning signals, some work has been done comparing individual versus combined sensory stimulation. In general, detection of bimodal signals is as good as, or better than, detection of signals stimulating only one sense[13,14,15,16,17], because the redundancy compensates for loss of attention in one of the senses. Furthermore, an auditory cue to indicate the quadrant in which a scope signal appeared has been used effectively to reduce visual search time and break up habitual search patterns[18].

In a somewhat different experiment on redundant information, Sumby and Pollack[19] found that speech in noise was perceived better when the speaker's lips could be seen, although Decker and Pickett[20] found that a meter indication of a speech signal in noise failed to increase the intelligibility. Tolhurst[21] showed the importance of integration of audio and visual cues by demonstrating a slowing of performance in speakers repeating messages when there was a lack of synchrony between lips and voice of the reader from whom the messages were initially received.

These foregoing combinations of signals generally required very little integration of information, either because the information content was small in at least one message, or because of high redundancy between messages. The field of audio-visual

aids to education has produced some studies of integration of more complex material, but with less control of the information in the material. A program at the University of Pennsylvania to study motion pictures as training aids produced several reports[22,23,24] that demonstrated the advantages of the sound motion picture film over the silent in training. The advantage of the sound track varied considerably with the nature of the material to be learned, the nature and amount of sound used, and the other situational factors. An earlier study[25] had demonstrated greater effectiveness in advertising with an audio-visual presentation than with audio or visual alone. The acceptance of this finding is demonstrated daily on commercial television.

A few direct attempts to study the unburdening of one sense by spreading the incoming information over two senses have had disappointing results. Although Flybar (flying by auditory reference) experiments suggested that complex auditory displays could be used in place of certain visual displays in aircraft[26] a number of inherent problems in auditory space perception have prevented their adoption.

Mowbray[27] presented subjects with simultaneous pairs of messages (one aural, one visual) that provided the necessary information for entering details on a map. His subjects made fewer correct responses than could be predicted for pure guessing, and (when uninstructed) were not even aware of the simultaneous presentation of data. He concluded that "... complex perceptions involving language cannot be effected by different sensory modes at the same time," conceding, however "... that very brief symbolic material, such as letters and numerals or simple words, may be handled effectively by two channels at the same time when presentation is precisely synchronized"[27] (p. 92).

Goldman[28] tested manual, compensatory tracking with a joystick and with a pair of linear controls (one for each hand) for a target varying in two dimensions. With each type of control, he compared performance with an all-visual display of the target on a CRT and performance with a bimodal display, presenting the horizontal dimension of target variation on the scope visually and the vertical dimension aurally, using an auditory pitch scale. Tracking with the all-visual display was consistently superior to tracking with the bimodal display at three levels of task difficulty.

What is wrong, then, with the concept of relieving the burden on one sense by putting another sense to work? Fitts[29] (p. 1314) suggested that human perception may be essentially a one-channel system, and that using two input channels may require an alternation that actually adds a burden to the information processing system rather than relieving it. In the case of competing messages in audition, Broadbent[12] found evidence of sequential processing as did Davis[30] with alternating auditory and visual stimuli. Kristofferson[31] has recently shown evidence of a basic time quantum, the same for different sensory channels, which he believes establishes a periodicity that controls the information processing stages of reaction time and the timing of the switching of attention among channels.

Adams and Creamer[32] tested the hypothesis that the single central channel is used primarily to reduce uncertainty as to what is going to happen rather than when it will happen. They had subjects respond with one hand to visual stimuli, with the other to auditory stimuli, simultaneously presented with uncertainty as to what the stimuli would be. Responses were poorer to bisensory conditions than to an all-visual control condition. They

concluded that the human operator is a one-channel system, providing the channel is concerned with resolving event uncertainty.

IMPLICATIONS OF THE LITERATURE

With a growing body of evidence in favor of a single-channel model of perception, why should we still be concerned with bimodal stimulation? First, our everyday experience shows that we frequently make more sense out of an environment that stimulates many senses at the same time than we do out of a single-sense situation, the various senses contributing various dimensions to our knowledge of the environment. Second, the single-channel concept is not definitely established, although there is strong experimental support for it. Third, single-channel processing does not necessarily rule out the utility of simultaneous sensory stimulation.

With regard to this third point, a model of sensory information processing proposed by Sperling[33] is of interest. Sperling assumes a central data processing channel, with separate short-term memory or storage capabilities associated with different senses, permitting one of a pair of competing messages to be stored temporarily while waiting its turn in the processing channel. He summarizes experimental data from a number of studies that demonstrate considerable differences between the temporal characteristics of visual and auditory information storage, with a faster fading of visual imagery. He then models a rapid scanning of visual data before fading, a conversion to auditory information storage through recoding, and a refreshing of the memory through a rehearsal loop (as exemplified by looking up a number in a telephone directory but repeating it orally to remember it before dialing). The importance of Sperling's model for

this discussion is his evidence of differences in storage characteristics for different senses. Thus, even if a single-channel model demolishes the unburdening argument for multimodal stimulation, such stimulation may still be feasible and useful if the messages are arranged and assigned to senses in such a way as to exploit these short-term storage characteristics. For example, selection of input channels so as to minimize the necessity for recoding and rehearsal might be of value in a display situation.

In summary, then, simultaneous presentation of highly redundant material is effective, especially in warning situations. Audio-visual training films are more effective than silent films. Attempts to unburden an overloaded sense by intersensory sharing of different information, on the other hand, have been notably unsuccessful. Yet everyday experience suggests that the human does respond intelligently to an integrated perception of multi-sensory stimulation. The potential advantages of exploiting this natural characteristic of man in the display situation still seem to outweigh the arguments against an unburdening effect and the negative results of a few experiments in warranting further consideration of factors involved in multimodal presentation of information.

An exploratory, experimental look at the implications of bi-sensory integration for command and control centers was, therefore, undertaken. The primary objective was to see whether or not two messages received at the same time via different senses could be usefully integrated. A second objective was to see whether there was any demonstrable advantage to using two senses instead of one. It was also hoped that the results of our experiments might give hints as to the importance of such parameters as task difficulty and temporal factors and even possibly shed additional light on the one-channel model of perception. In line with these objectives, two experiments were performed.

SECTION III

EXPERIMENT I

DESIGN OF EXPERIMENT I

Purpose: Experiment I was intended to be exploratory. The first objective was to attempt a demonstration that two verbal messages, received through different senses, can be usefully integrated. The effects of message difficulty on integration were also of interest.

Experimental Task: Vision and audition were selected as the senses to be tested, since both serve regularly in everyday life as channels for the input of verbal information.

The basic format selected for the experiment was similar to that of standard reading comprehension tests - that is, a presentation of information for a controlled period of time followed by a question on that information. In our case, however, the information was presented in the form of two messages: one presented visually, one presented aurally. The question required information from both messages to select the correct answer from four alternatives. To avoid bias in favor of certain of our college-student subjects, we drew our stimulus material from a wide range of subject matter, including history, psychology, sociology, geology, medicine, simple arithmetic, and miscellaneous other areas. Elementary text books, popular nonfiction, health leaflets, and other sources were searched for simple declarative sentences encompassing enough different but related ideas to permit the construction of pairs of messages

that could be integrated to answer a question on contents. Nearly all source material required rewriting before it was suitable for test items, and many other items were wholly originated by the experimenters.

An attempt was made to vary difficulty by varying the length of the messages and the number of separate concepts or ideas in each message. Quantitative scaling for difficulty was not attempted beyond a word count for each statement and a count of arbitrarily defined "ideas" in each statement, as agreed upon by the author and his assistant.

The experiment consisted of fifty items, each made up of two messages and a four-choice question. Individual messages varied in length from four to thirty words: the difference between numbers of words in message pairs varied from zero to sixteen. Individual messages varied in number of ideas from one to four ideas; the difference between numbers of ideas in message pairs varied from zero to three. Insofar as was feasible, word counts and idea counts were balanced between Messages 1 and 2. The tasks inherent in the questions included recall of two or more facts, reasoning based on two or more facts, and arithmetic combination of two or more quantities. Table I gives examples of test items.

The test items were prepared in two forms. In Form A, Message 1 was always presented aurally, while Message 2 was presented visually. In Form B, the modes of presentation were reversed: Message 1, visual; Message 2, aural. By combining the visual presentations of Forms A and B, Form C, all visual, was formed as a control condition.

Visual presentation was accomplished by preparing a 35 mm slide for each message. These slides were projected by a Kodak Carousel 35 mm slide projector with a "zoom" lens. In Form C,

TABLE I

EXAMPLES OF TEST ITEMS

EXAMPLE 1

Message 1. The earth was first a molten mass.

Message 2. The earth was originally enveloped in a cloud of hot vapor.

Question: In the beginning the earth was

a) a molten mass, directly exposed to the sun.

Answer: b) a molten mass, surrounded by a cloud of hot vapor.

c) a solid mass, directly exposed to the sun.

d) a solid mass, surrounded by a cloud of hot vapor.

	Message 1	Message 2	Difference
Word Count:	7	11	4
Idea Count:	1	1	0

Task: Recall

TABLE I (continued)

EXAMPLE 2

Message 1. The letter on the badge is the initial of the home state.

Message 2. Blue badges are from the north, gray badges are from the south.

Question: The wearer of a gray badge with the letter "M" is from

a) Montgomery, Alabama

b) Madison, Maine

c) Milton, Massachusetts

Answer: d) Biloxi, Mississippi

	Message 1	Message 2	Difference
Word Count:	12	12	0
Idea Count:	1	2	1

Task: Reasoning

TABLE I (continued)

EXAMPLE 3

Message 1. Normally the drive downtown takes one hour.

Message 2. Each of the following adds 15 minutes to the drive: rain, rush hour, darkness.

Question: How long does it take to drive downtown at midnight in the rain?

Answer:
a) 75 minutes
b) 90 minutes
c) 105 minutes
d) 120 minutes

	Message 1	Message 2	Difference
Word Count:	7	14	7
Idea Count:	1	4	3

Task: Arithmetic

the all-visual control study, pairs of slides were presented together, one above the other, using two projectors operated manually by one experimenter. All questions were also projected via slides. Aural presentation was accomplished via magnetic tape played through a single-channel Revere tape recorder. The same voice was used in the recording of all tapes. A Kodak Carousel Programmer permitted a signal to be recorded on the tape that would automatically operate the slide projector, thus synchronizing the slides with the voice record. (In practice, difficulties with the tape recorder forced us to run two series by reading aloud from a script while manually operating the slide projector. The group results from these series did not differ significantly from the results of the automatically controlled series, according to "t" tests; so these conditions were not differentiated in the treatment of results.)

There was little in the literature to guide us on the selection of presentation times. In order to avoid confounding the effects of time and item length, and to standardize experimental procedure, we decided to hold presentation time and decision time constant for all items. Timing of the reading of items and preliminary experiments led us to select 15 seconds for presentation of the messages, followed by 20 seconds for answering the question, or 35 seconds for each item. The longest messages could be read aloud intelligibly in just under 15 seconds; shorter messages were placed near the center of the presentation period by introducing a measured time delay between the slide-change signal and the beginning of the reading when we recorded the tapes.

One form of the test, together with instructions, practice items, and rest breaks, could just be comfortably run in a normal class period. In this first study, we elected to run our test

and control conditions on separate groups of subjects, thus introducing the possibility of error due to individual differences between groups, rather than to have subjects serve as their own controls, thus confounding the results with learning effects. Using separate groups also permitted us to study the results of the major variables before deciding what control studies might be required. Thus Form C, an all-visual control experiment, was run on only those ten items of Forms A and B that yielded enough errors to merit analysis.

Subjects: The subjects of Experiment I were 76 undergraduate students at Tufts University, 31 male, 45 female. They all volunteered to take part and were paid for their services. Form A was administered to 37 subjects, Form B to 39. Form C, a control series, was administered to 19 students, fifteen male and four female, at a later time.

Procedure: All testing of Forms A and B was conducted in the same classroom at Tufts University on three days in the same week, for a total of 7 testing sessions. With room lights off, window lighting left the room dim enough for good contrast on the projection screen, yet light enough for subjects to see and write on their answer sheets. Form C was run in two sessions some three months later in a different room -- the second session requiring overhead lighting. (The mean results between the two Form C groups were not significantly different and were treated as a single set in further analysis.)

Prior to testing, the projector, programmer, and tape recorder were installed, checked, and warmed up. The projector was focused and "zoomed" so that the horizontal dimension of the slides filled the screen. An answer sheet containing item numbers for ten preliminary and fifty regular items, with the letters: "a, b, c, d" following each number, was placed at each seat.

When all subjects had arrived and were seated, they were briefly instructed on entering identification data on the answer sheet and on marking answers. Formal instructions were then presented via the tape recorder, with a break for questions, followed by ten practice trials of increasing difficulty. After the practice, a short rest period permitted additional questions to be answered. These preliminary activities took about ten minutes.

The main experiment followed, with a continuous sequence of 15 seconds of message presentation followed by 20 seconds to answer the question appearing on the screen. When any question had been on for 20 seconds, it was immediately followed by the visual message of the next item. As mentioned above, there was a delay between the beginning of a slide message and the beginning of the taped message depending on the length of the taped message, so that each aural message was read during the middle portion of the slide presentation. After the first twenty-five items there was a rest period of two or three minutes followed by the second twenty-five items, concluding the series.

RESULTS OF EXPERIMENT I

Group Results: The mean error scores for the group and for certain subgroups are presented in Table II. The outstanding characteristic of the results of Experiment I is the low error rate. Seventy-six subjects each answered 50 questions, for a total of 3800 answers. Only 433 answers were incorrect, giving an overall error rate of 11.4 percent and an average of 5.7 errors per subject. The distribution of errors is given in Figure 1. The worst score was 15 errors, made by one subject; two others had 14 errors and 11 errors was the next worst score. Thus in the worst case, 70 percent of the items were answered correctly,

TABLE II

SUMMARY OF MEAN ERROR SCORES, EXPERIMENT I

Group	N	Mean	S.D.	Difference	"t"
Total	76	5.70	2.96		
Form A	37	5.46	2.85	A-B 0.46	A-B 0.67*
Form B	39	5.92	3.06		
Male	31	5.90	2.89	M-F 0.35	M-F 0.51*
Female	45	5.55	3.09		

*Not statistically significant.

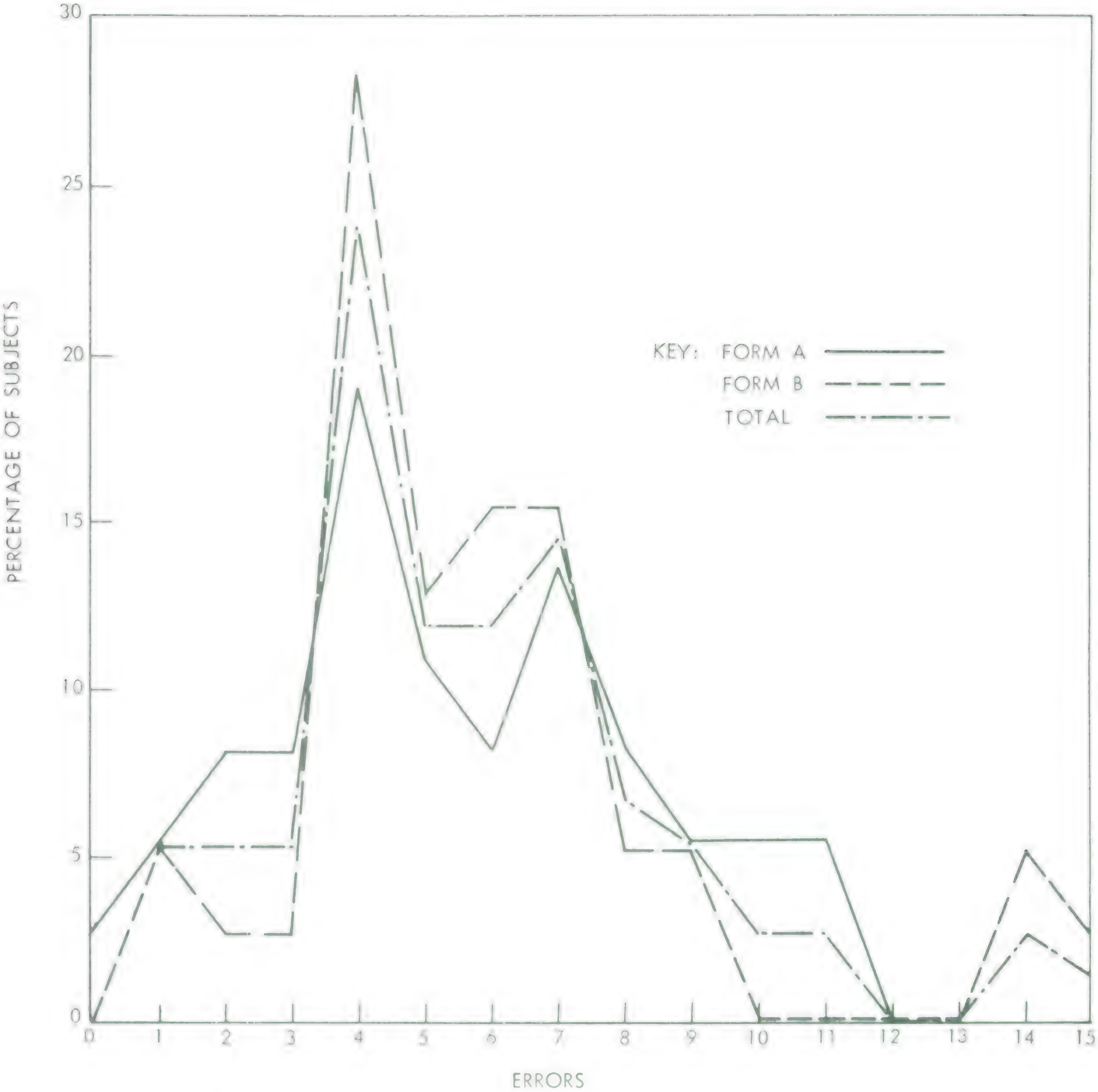

Figure 1a. Distribution of Errors, Experiment I - Form A versus Form B.

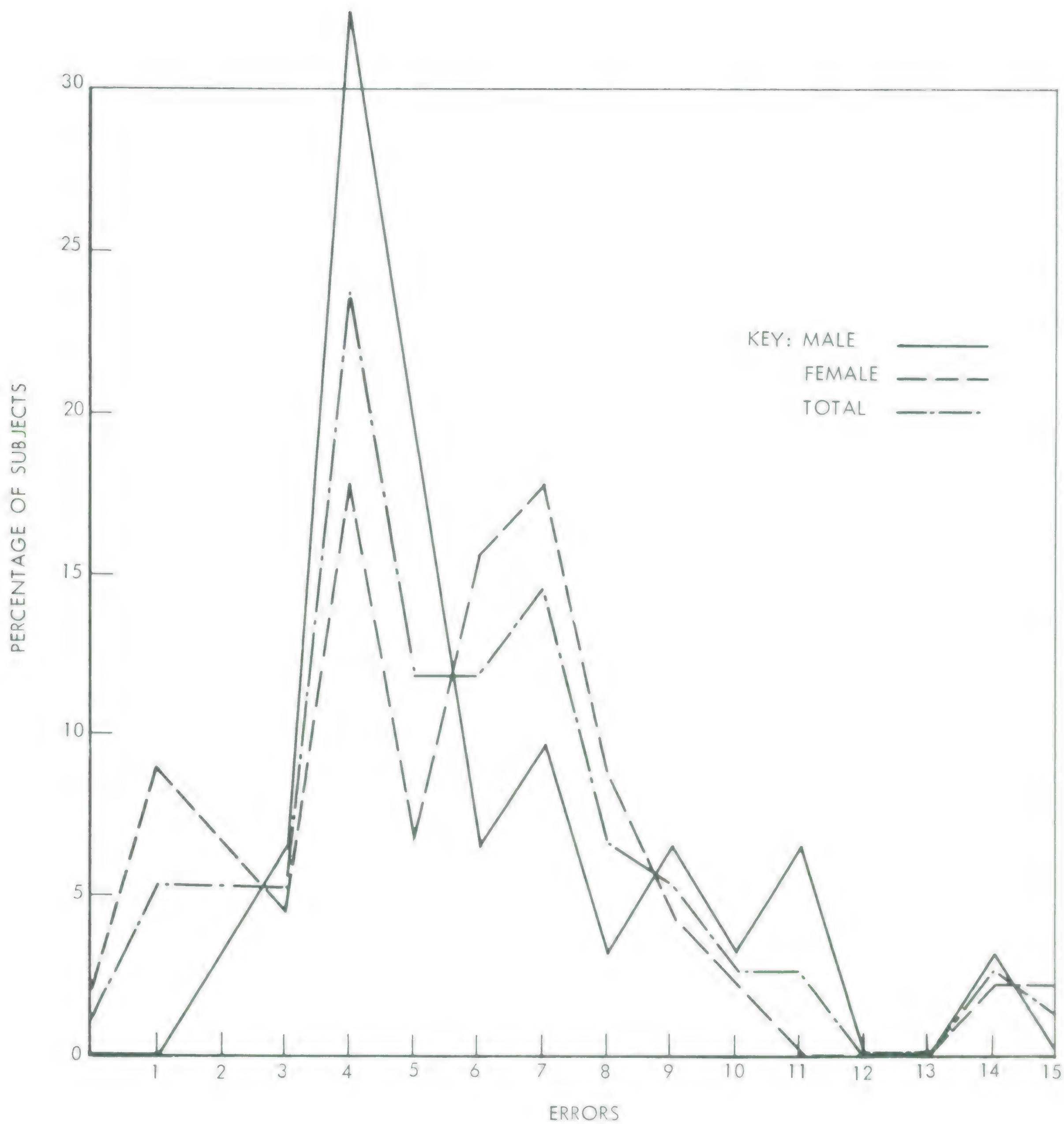

Figure 1b. Distribution of Errors, Experiment I - Males versus Females.

and on the average, 89 percent were correct. One subject made a perfect score, all items correct; four others had only one error, while still four more made two errors. Obviously, under the conditions of the experiment, subjects could and did integrate the contents of two items presented via different senses.

A comparison of the results of Form A and Form B (See Table II) shows no statistical significance according to the "t" test. Because of considerable departure from normality, the error distributions for the two forms are shown graphically in Figure 1a. It is concluded from their similarity that the assignment of messages between senses in constructing the two forms did not introduce any bias in favor of one form over the other.

Although a "t" test yielded no significant difference between the mean scores of male and female subjects, the distributions of their scores (See Figure 1b) suggests that the secondary mode of the group at 7 errors was contributed mainly by the female subjects, while the preponderance of females at 0 and 1 errors pulled their mean score back below that of the males. The nature of the distributions and the numbers of subjects involved do not permit any profound conclusions to be drawn from these observations.

The difficulty of test items was hypothesized as a function of the number of words and the number of ideas in the item. Figure 2 plots scattergrams of errors versus number of words and number of ideas. There is no tendency evident for positive correlation of errors with these two parameters. In fact, we can see a slight tendency toward a negative correlation between words and errors for the ten most difficult items.

Although the low error rate in our results makes further statistical analysis fruitless, additional information may be gained

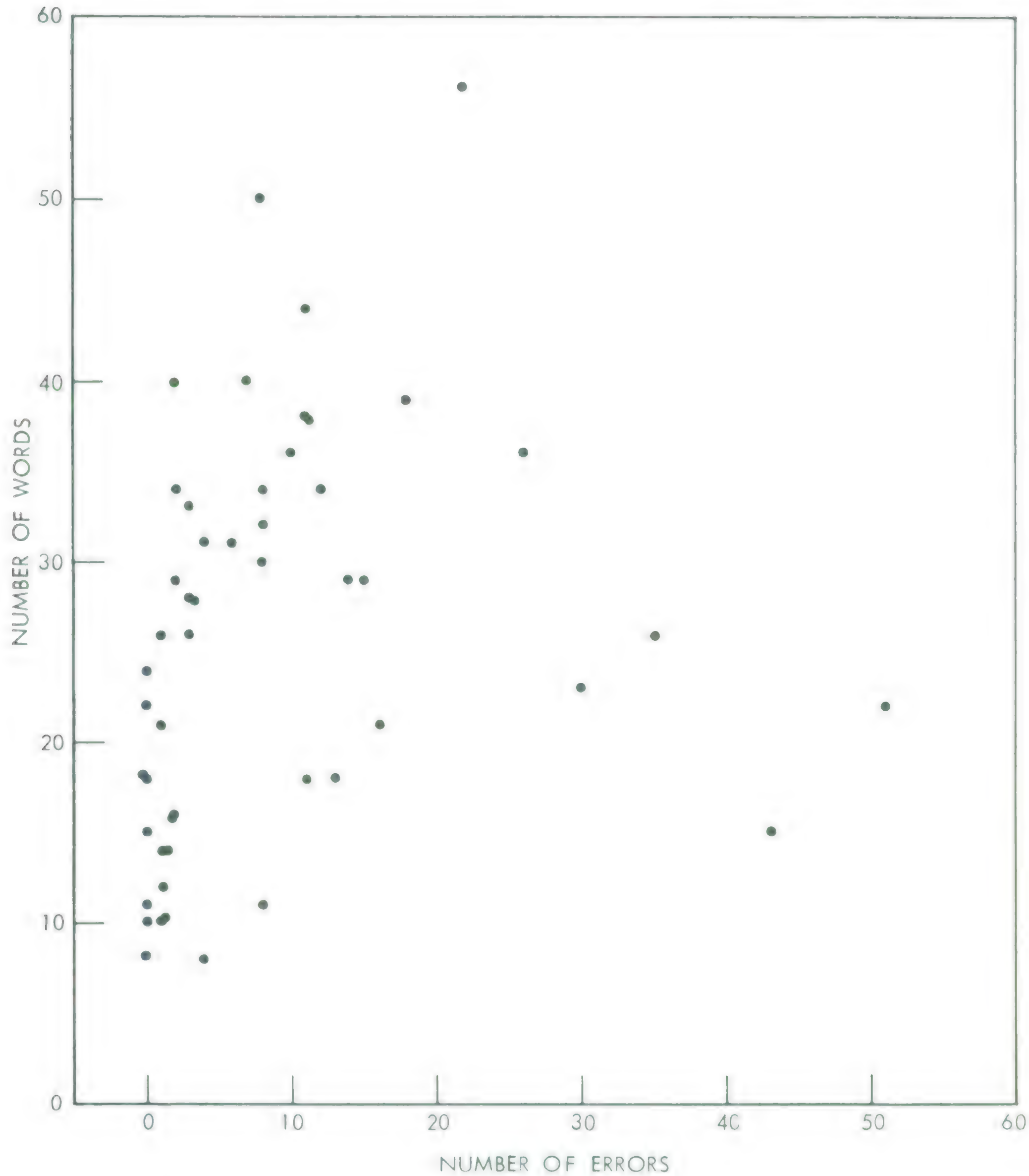

Figure 2a. Correlation of Errors versus Item Characteristics, Experiment I - Errors versus Number of Words.

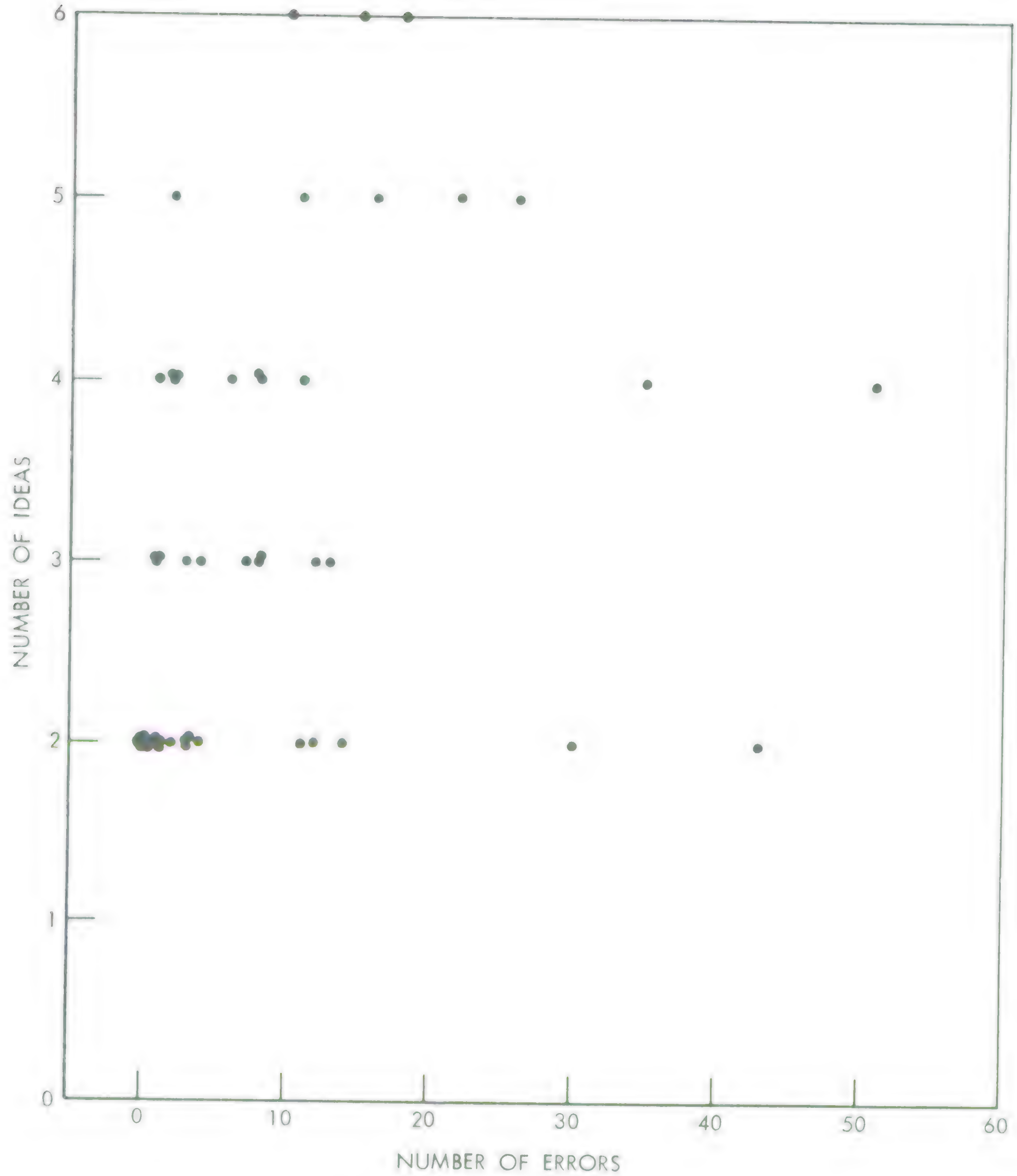

Figure 2b. Correlation of Errors versus Item Characteristics, Experiment 1 - Errors versus Number of Ideas.

by examining the individual test items. The items have been arranged in rank order of decreasing difficulty -- that is, rank 1 has been given the item on which the greatest number of errors occurred, rank 50 the item yielding the fewest errors. Figure 3 plots cumulative error as a function of error rank. We can readily see that 50 percent of all errors were contributed by seven items, while the most difficult half of the test accounted for over 90 percent of the errors. At the other end of the curve, we note that 8 items contributed no errors. In fact, another 8 items contributed only 1 error each, and still 5 more contributed only 2 errors each.

Although no significant difference was found between Forms A and B with regard to group performance, certain individual items tended to show a difference between forms, suggesting that an interaction occurred between messages and the sense through which the message was presented. A search for items in which there was a difference of 5 or more people between those making errors on Form A and those making errors on Form B yielded three items ... the fifth, sixth, and ninth items in order of total difficulty. We shall return to these items later.

<u>Control Study</u>: Obviously, most of the test was too easy to yield meaningful data. So the ten most difficult items, including over 60 percent of the errors and the three items showing differences greater than 5 between forms, were selected for more detailed analysis and for a control experiment.

The control experiment was run some three months after Forms A and B. It occupied the first half of a class period, the second half being devoted to a pilot run of Experiment II. Instructions and procedures were similar to those of the earlier sessions, except that all pairs of messages were presented together visually,

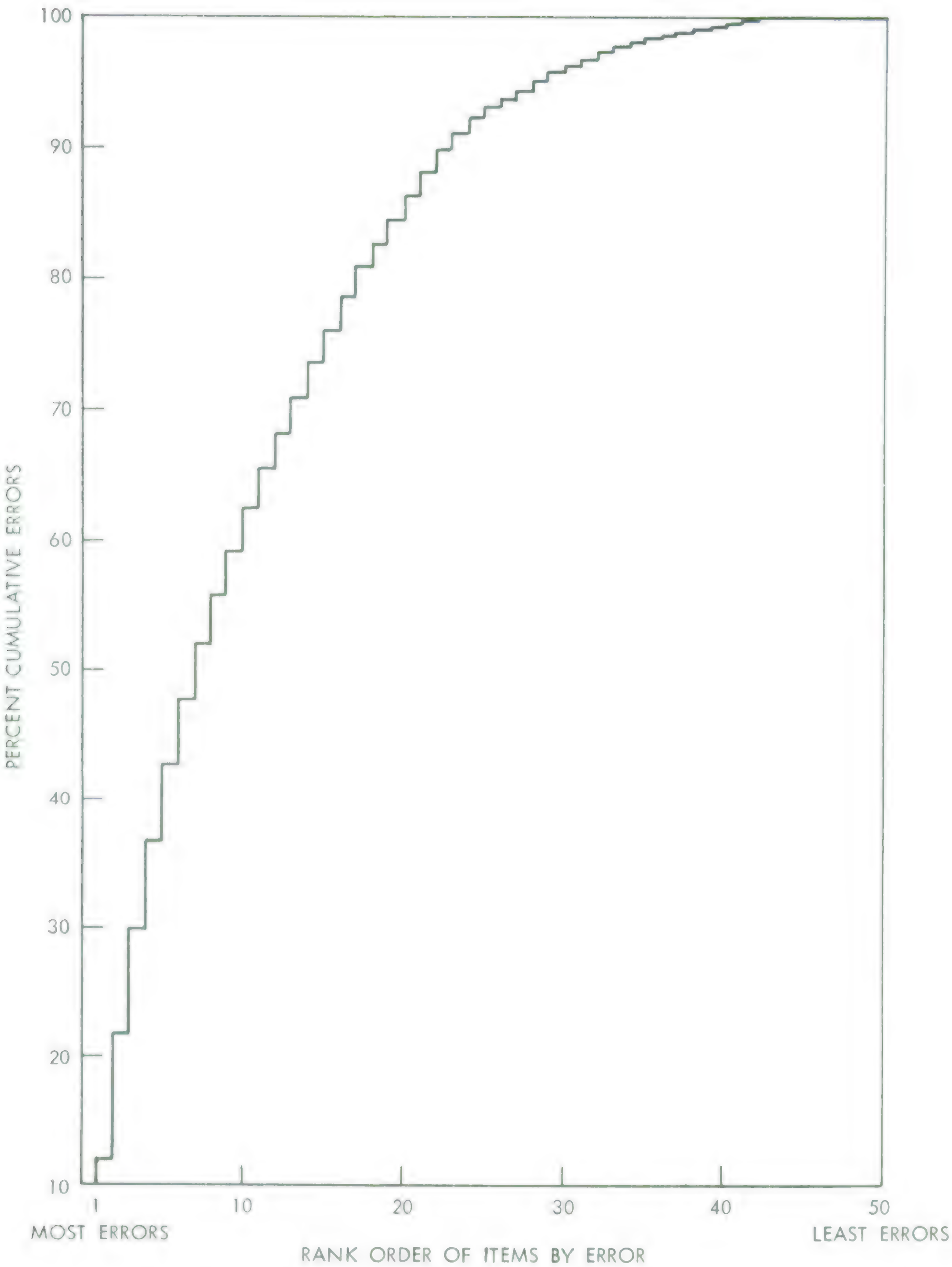

Figure 3. Cumulative Distribution of Errors by Item, Experiment I.

one above the other via two projectors. The 15-second presentation time and 20-second decision time were preserved. Practice was given on five preliminary items taken from the error-free items of the main experiment. From here on, the results of this all-visual control experiment will be reported as Form C and compared with the results of the same ten items of Forms A and B.

Table III summarizes the group error scores and Figure 4 compares the distributions of errors in the control study and the combined data from the ten most difficult items of Forms A and B. Although the "t" test of the mean difference between (A+B) and (C) falls far short of statistical significance, the irregularity of the distributions does not permit us to accept a "no difference" conclusion with confidence. If anything, there is a slight tendency toward fewer errors in the all-visual control situation. There is certainly no evidence of any advantage in the bimodal presentation.

Although there was some shifting of order of difficulty among the ten items between the bimodal and all-visual presentations, the statistically significant rank-order correlation between the two sets of results permits us to accept the hypothesis that the order of difficulty was basically the same in both cases.

The evidence presented so far, then, leads us to conclude that, as a group, the ten most difficult items in Forms A and B were not difficult because of the bimodal presentation, since essentially the same error rate was yielded in the all-visual control study. Because our scattergrams (See Figure 2) failed to demonstrate any relationship between difficulty and simple counts of words and ideas, we conclude that the <u>tasks</u> required of the subject were particularly difficult in these ten items. Since we are not concerned here with the nature of difficult

TABLE III

SUMMARY OF ERROR SCORES, TEN DIFFICULT ITEMS OF FORM (A + B) AND FORM C (ALL-VISUAL CONTROL), EXPERIMENT I.

Form (A + B), N = 76		Form C, N = 19	
Item Rank	Errors	Errors	Item Rank
1	51	11	1
2	43	10	2
3	35	6	5.5
4	30	7	4
5	26	8	3
6	22	6	5.5
7	18	3	9.5
8	16	4	7.5
9	15	3	9.5
10	14	4	7.5
Mean Error	3.55	3.26	
S.D.	1.62	1.71	
Difference	0.29		
"t"	0.65 (not statistically significant)		

Rank-Order Correlation (A + B) vs. C. $\rho = 0.86$

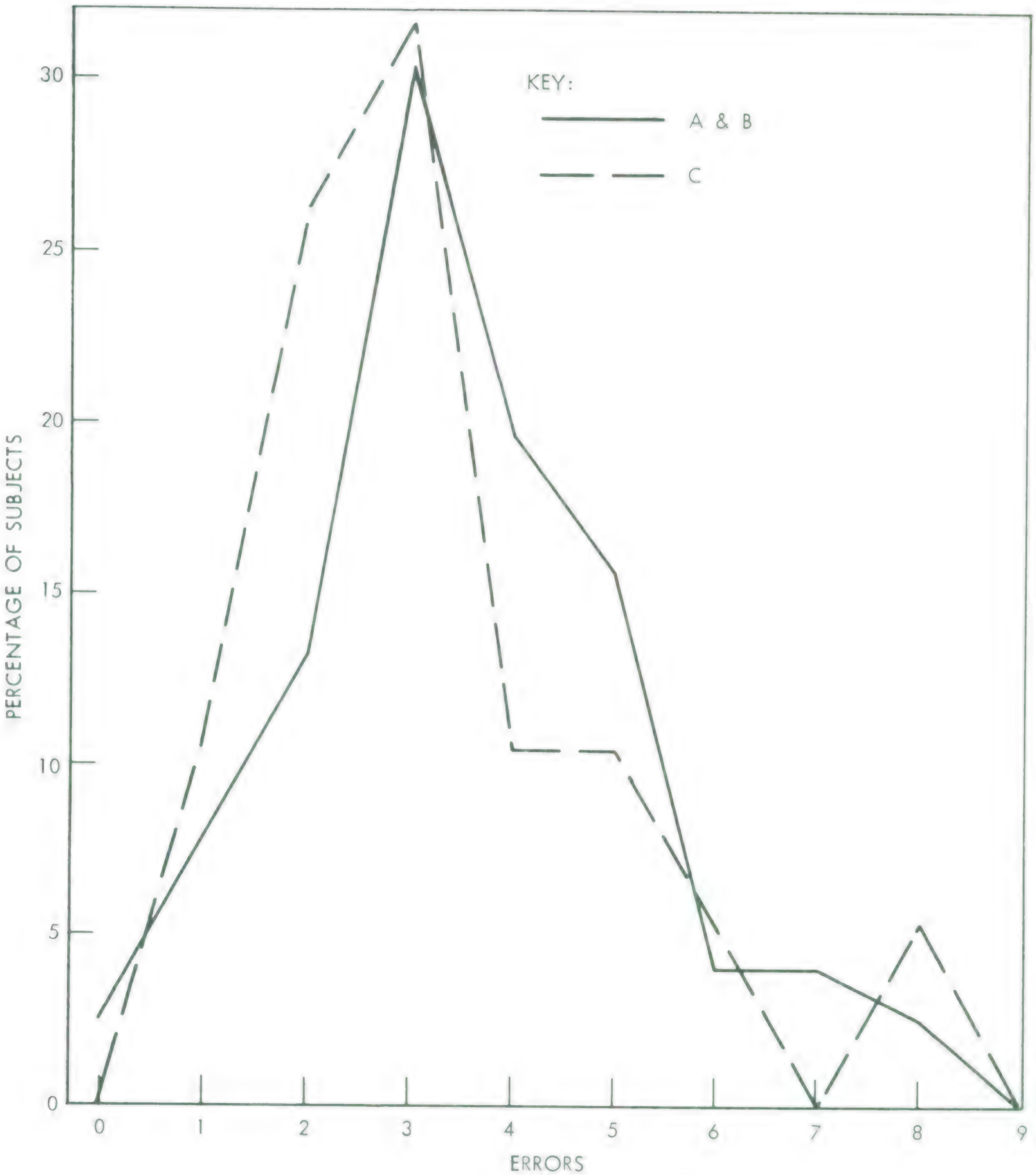

Figure 4. Error Distribution for Worst 10 Items, Experiment I.

items *per se*, we will not analyze all ten items individually. However, we will now return to those three items that differed markedly in error scores between Forms A and B, thus hinting at a possible relationship with mode of presentation.

Item Analysis: Tables IV, V, VI present the three items in question, together with a summary of the number of subjects selecting each alternative. The item showing the greatest difference between Forms A and B (See Table IV) was sixth most difficult. In both Forms A and B, the most favored wrong answer was c, which was selected much more frequently in Form B. The key word that permits differentiation between the correct answer, a, and c, is "prenatal," appearing in the middle of Message 1. So, the key word was missed most frequently when it was presented visually. One hypothesis to explain this occurrence might be that when both vision and audition are being stimulated audition tends to dominate. An alternative explanation could be that there was time to alternate attention between senses. The visual presentation appeared first. We can imagine a switching of attention when the new message started, with a return to the visual message following the auditory. Since Message 2 is long, the silent periods before and after its auditory presentation were short; thus many subjects may not have read as far as the key word either before or after the auditory presentation. Such an analysis is consistent with the selection of answer b by one subject taking Form A, but some further reason, probably simply the missing of the significance of the word "prenatal," must have lain behind the five selections of c in Form A. The all-visual presentation yielded three selections each of b and c, suggesting that the key words of either message were equally likely to be missed under that condition.

TABLE IV

TEST ITEM WITH GREATEST DIFFERENCE BETWEEN ERROR SCORES OF FORM A AND FORM B. (SIXTH MOST DIFFICULT ITEM, EXPERIMENT I.)

Message 1. The 20 primary ("baby") teeth begin to form around the second month of prenatal life, and by birth a considerable part of the crowns is already formed.

Message 2. The primary teeth erupt and are shed at varying times. Although often called "baby" teeth, some remain in the child's mouth until he is 11 or 12 years old.

Question: Primary teeth begin to form

a) several months before birth, and some last until age 11 or 12.

b) several months before birth, and are all lost by age 8.

c) several months after birth, and some last until age 11 or 12.

d) several months after birth, and are all lost by age 8.

Number of Responses

		Form A	Form B	Form C
Message 1		Aud.	Vis.	Vis.
Message 2		Vis.	Aud.	Vis.
Correct	a)	31	23	13
	b)	1	0	3
	c)	5	14	3
	d)	0	2	0

TABLE V

TEST ITEM WITH SECOND GREATEST DIFFERENCE BETWEEN ERROR SCORES OF FORM A AND FORM B. (NINTH MOST DIFFICULT ITEM, EXPERIMENT I.)

Message 1. John always uses a gun, Bill a knife, Bob a blackjack, and Ray a club.

Message 2. The police found a knife and a club at the scene of the crime.

Question: The most logical suspects were

a) Bill and Ray

b) Bob and Ray

c) John and Bob

d) John and Bill

Number of Responses

		Form A	Form B	Form C
Message 1		Aud.	Vis.	Vis.
Message 2		Vis.	Aud.	Vis.
Correct	a)	25	36	16
	b)	11	1	0
	c)	1	2	2
	d)	0	0	1

TABLE VI

TEST ITEM WITH THIRD GREATEST DIFFERENCE BETWEEN ERROR SCORES OF FORM A AND FORM B. (FIFTH MOST DIFFICULT ITEM, EXPERIMENT I.)

Message 1. In spite of torrential early rains, the oceans at first contained fresh water, because there was no soil to erode.

Message 2. The gradual erosion of rocks over a long period of time changed the seas to brine.

Question: Our seas became salty through

a) rock erosion when the topsoil washed away.

b) rock erosion, because there was no soil.

c) condensation of mineral-laden clouds.

d) flooding of soil-laden valleys.

Number of Responses

		Form A	Form B	Form C
Message 1		Vis.	Aud.	Vis.
Message 2		Aud.	Vis.	Vis.
	a)	7	16	7
Correct	b)	28	22	11
	c)	1	0	0
	d)	1	0	1

The item of Table V was the ninth most difficult. Answer b was the favored wrong answer, but only in Form A. The task requires associations rather than memory of key words, and the errors occurred when most of the associations (Message 1) were presented aurally, thus not supporting the hypothesis of auditory domination. Since both messages were relatively short, there was time for reference to the visual message after the auditory presentation. Such alternation between messages should result in fewer errors when the message containing the most associations to be memorized was the one available for a second reference, as was the case. The results of the all-visual presentation are consistent with the alternation hypothesis, too.

The item of Table VI was the fifth most difficult. Here the key words differentiating between the correct answer, b, and the most favored wrong answer, a, were "no soil," appearing near the end of Message 1. This error occurred most frequently when the key words were presented visually. The analysis is almost identical to that of the item in Table IV, being consistent with either an auditory domination or a sensory alternation hypothesis. However, the seven selections of a in the all-visual presentation are more difficult to explain and may indicate an inherent ambiguity in the item.

CONCLUSIONS FROM EXPERIMENT I

Experiment I was too easy to yield conclusive results. Certainly, it demonstrated that, with 15 seconds of presentation and 20 seconds for decision, bimodal presentation of short messages permits integration of the information within those messages, although it has no advantage over presentation via a single mode. Whether there was simultaneous processing of information

in two channels or alternation between the two senses cannot be determined, although there is a little evidence that suggests alternation. Varying message length turned out to be an inadequate technique for varying task difficulty in this experiment.

SECTION IV

EXPERIMENT II

DESIGN OF EXPERIMENT II

Purpose: Experiment II was a second attempt to demonstrate the integration of two verbal messages received through different senses, and to compare this mode of presentation with all-visual presentation of the same messages. A further objective was to demonstrate the applicability of bimodal presentation to Air Force operational problems by basing the experiment on a simulation of an Air Force function.

Experimental Task: Again, the task had to require a decision that could be made only by combining information received through two senses at the same time. Because of the possibility that attention was alternated between the senses in Experiment I, it was considered desirable to fill the presentation time completely with the aural presentation. The results of Experiment I suggested that the messages might as well all be of the same length, since variable length did not contribute a difficulty dimension to the study. So a message format was sought in which all messages contained the same number of words. As a means of minimizing alternation of attention between senses, it was decided to reduce the difference between the time required for silent reading of a visual message and the time required for the aural presentation of a message by using very short messages. Finally, a task had to be devised that had some relevance to Air Force problems.

The task finally devised was a game or exercise simulating the aircraft identification function of an air defense center. The simulation was admittedly crude, since the criteria for item length, bimodal presentation, a decision requiring both messages, and decision time all took precedence over realism.

The task involved two messages, each of which contained two bits of information about an aircraft. The combined information, together with a set of ground rules, permitted the subject to determine one of four alternative answers.

One message gave aircraft position in terms of two zones -- Alpha and Bravo -- and aircraft altitude (Angels) as either High or Low. The other message told whether or not an identification signal was being emitted by the aircraft -- ID YES or ID NO -- and whether the aircraft type was single-engine or multi-engine -- Type Single or Type Multi. A typical pair of messages, then, was:

Message 1: Zone Alpha Angels High

Message 2: ID No Type Multi

There was, therefore, a vocabulary of 16 pairs of messages, all messages of equal length, even to the number of syllables. The choices for identification were always the same:

a. Military

b. Civilian

c. Hostile

d. Unknown

Several ground rules were devised such that an aircraft could either be definitely called Military, Civilian, or Hostile or, in the absence of enough information for such a choice, Unknown. The rules were also devised so that both messages were required for a decision, although either message could contain superfluous information. The seven rules follow:

1. "ID Yes" is either Military or Civilian.

2. "ID No" is either Hostile or Civilian.

3. There are never any High, Single, Civilian aircraft.

4. There are never any Low, Multi, Military aircraft.

5. There are never any High, Civilian aircraft in Zone Alpha.

6. There are never any Low, Multi, Hostile aircraft in Zone Alpha.

7. There are never any Single, Civilian aircraft in Zone Bravo.

Some mnemonic aids were formulated in the form of a rationale for the rules (such as: Zone Bravo is over the ocean, and little single-engine private planes cannot fly out that far -- Rule 7).

Three series of 48 items were prepared. Each series was formed by cycling the 16 items three times in a modified random order. In Series A, the message pairs appeared together on a single slide for each item, thus constituting an all-visual control series. In Series B, the ID-Type messages were on slides, while the Zone-Angels messages were recorded on magnetic tape. In Series C, the Zone-Angels messages were on slides, the ID-Type messages on tape.

Presentation time was selected (on the basis of pilot experiments) as approximately 2 seconds, with 8 seconds for decision time. Slide-changing signals were put on the voice tape by an electronic timer. In practice, automatic slide changing did not work consistently; so in most series the slides were changed manually, still cued by the taped signal. The same voice recorded all taped messages.

In this experiment, subjects acted as their own controls. That is, each subject was given all three series of stimuli. Sequences of series were varied among the subjects, with three sequences being used:

Sequence I. ABC

Sequence II. BCA

Sequence III. CAB

Thus, each subject answered 144 items. Each item took 10 seconds to administer; so actual testing time was 24 minutes, leaving 16 minutes for instructions, practice, and rest breaks in a 40-minute class period. A pilot test indicated that this timing was feasible.

Subjects: The subjects of Experiment II were obtained both at Tufts University and at L. G. Hanscom Field. The Tufts students volunteered their services and were paid. The Hanscom military subjects were assigned to the task; the civilians volunteered. Hanscom subjects were not paid for their participation. There were 10 Tufts subjects, 4 male and 6 female. There were 27 Hanscom subjects, 24 male and 3 female. The 3 female subjects and 4 of the male subjects were Northeastern University students employed at Hanscom Field. The remaining 20 male subjects were Air Force personnel, 10 officers and 10 airmen.

Procedure: The testing at Tufts University was all done in a lecture hall in which a dim overhead lighting could be so adjusted that the answer sheets were legible without introducing glare on the projection screen. At Hanscom Field, a small classroom was used. Enough bulbs were removed from overhead fluorescent fixtures so that, again, both the screen and the answer sheets could be seen clearly.

Prior to testing, the projector and tape recorder were installed, checked, and warmed up. The projector was focused and "zoomed" so that the horizontal dimension of a slide filled the screen. Answer sheets were placed at all seats.

When all subjects had arrived and were seated, they were given instructions that included the purpose of the experiment, the nature and details of the aircraft identification exercise, and several sample problems. A rationale was given for the seven ground rules as a mnemonic aid, and a condensed version of the rules was printed at the top of each answer sheet (see Figure 5).

The main experiment followed, with alternating presentation of messages for 2 seconds (either bimodal or all-visual, depending on the Series) and 8 seconds to make a decision and mark the answer sheet. The screen was blank during the decision periods. Between Series, brief instructions were given as to the mode of presentation of the following Series. Sequences of Series were varied between different subject groups in order to distribute learning effects across all Series. At the end of the experiment, subjects were urged to contribute comments on the backs of their answer sheets. The Hanscom subjects were asked three specific questions:

1. Rank the Series in order of difficulty.
2. Which Series was preferred?
3. In bimodal Series, did message processing seem to be simultaneous or sequential?

RESULTS OF EXPERIMENT II

Group Characteristics: The mean error scores for the group and for certain subgroups are presented in Table VII. Thirty-seven subjects each made 144 decisions for a total of 5328 answers. There were 1938 incorrect answers, giving an overall error rate of 36.3 percent and an average of 52.4 errors per subject. The distribution of errors is shown in Figure 6. The results show that Experiment II was more difficult than Experiment I.

PROJECT 141 ANSWER SHEET

Experiment II - ____ Series _________ Date ____________________

Name ______________________________

Rules

Never High Single Civ

Never High Civ ALPHA Never Low Multi Hos	BRAVO Never Single Civ

Never Low Multi Mil

ID YES .... MIL/CIV
ID NO .... HOS/CIV

1.	MIL	CIV	HOS	UNK	17.	MIL	CIV	HOS	UNK	33.	MIL	CIV	HOS	UNK
2.	MIL	CIV	HOS	UNK	18.	MIL	CIV	HOS	UNK	34.	MIL	CIV	HOS	UNK
3.	MIL	CIV	HOS	UNK	19.	MIL	CIV	HOS	UNK	35.	MIL	CIV	HOS	UNK
4.	MIL	CIV	HOS	UNK	20.	MIL	CIV	HOS	UNK	36.	MIL	CIV	HOS	UNK
5.	MIL	CIV	HOS	UNK	21.	MIL	CIV	HOS	UNK	37.	MIL	CIV	HOS	UNK
6.	MIL	CIV	HOS	UNK	22.	MIL	CIV	HOS	UNK	38.	MIL	CIV	HOS	UNK
7.	MIL	CIV	HOS	UNK	23.	MIL	CIV	HOS	UNK	39.	MIL	CIV	HOS	UNK
8.	MIL	CIV	HOS	UNK	24.	MIL	CIV	HOS	UNK	40.	MIL	CIV	HOS	UNK
9.	MIL	CIV	HOS	UNK	25.	MIL	CIV	HOS	UNK	41.	MIL	CIV	HOS	UNK
10.	MIL	CIV	HOS	UNK	26.	MIL	CIV	HOS	UNK	42.	MIL	CIV	HOS	UNK
11.	MIL	CIV	HOS	UNK	27.	MIL	CIV	HOS	UNK	43.	MIL	CIV	HOS	UNK
12.	MIL	CIV	HOS	UNK	28.	MIL	CIV	HOS	UNK	44.	MIL	CIV	HOS	UNK
13.	MIL	CIV	HOS	UNK	29.	MIL	CIV	HOS	UNK	45.	MIL	CIV	HOS	UNK
14.	MIL	CIV	HOS	UNK	30.	MIL	CIV	HOS	UNK	46.	MIL	CIV	HOS	UNK
15.	MIL	CIV	HOS	UNK	31.	MIL	CIV	HOS	UNK	47.	MIL	CIV	HOS	UNK
16.	MIL	CIV	HOS	UNK	32.	MIL	CIV	HOS	UNK	48.	MIL	CIV	HOS	UNK

Figure 5. Sample Data Sheet, Experiment II.

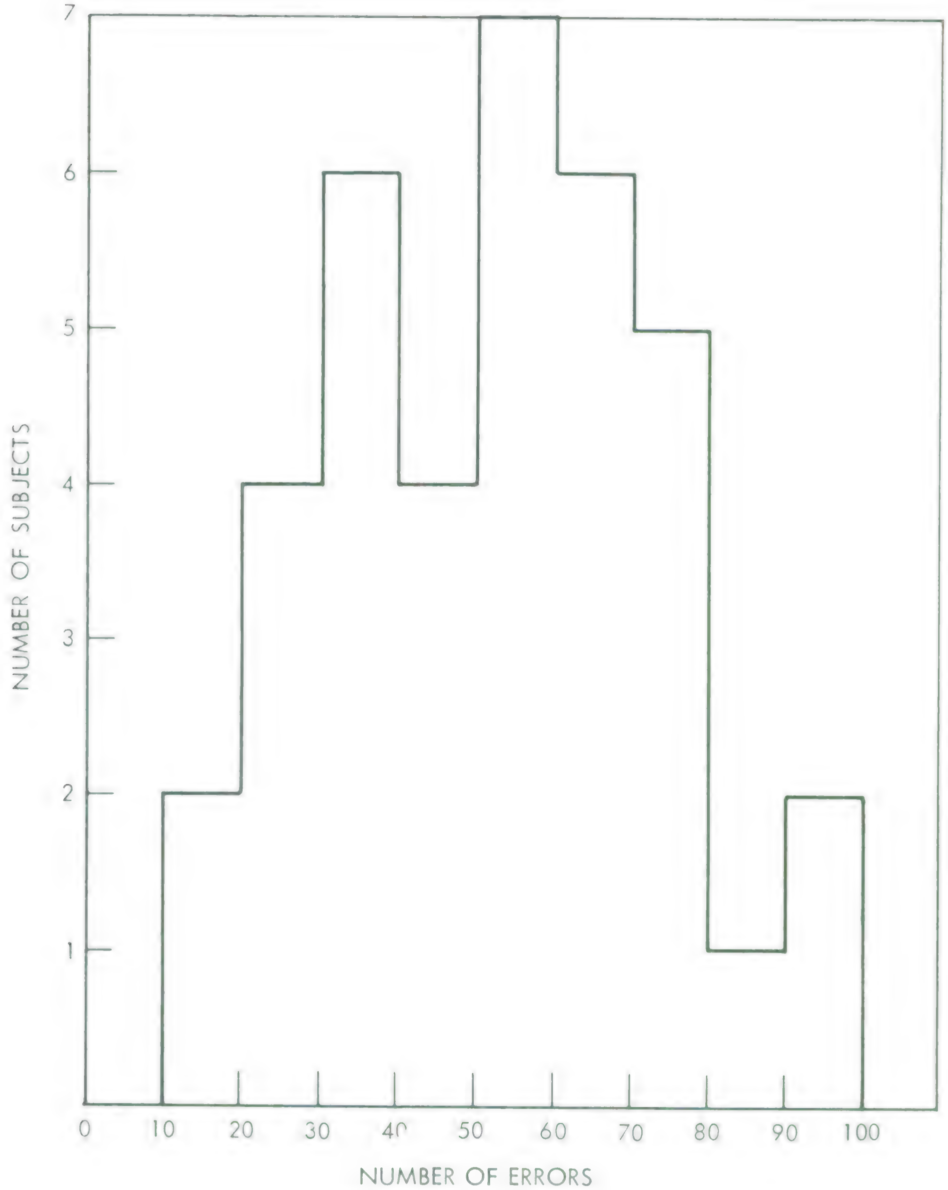

Figure 6. Distribution of Error Scores, Experiment II.

TABLE VII

SUMMARY OF MEAN ERROR SCORES, EXPERIMENT II

Group	N	Mean	S.D.
Total	37	52.4	21.4
Sub-Group			
Tufts Subjects	10	45.3	
Hanscom Subjects	27	55.0	
Students	17	45.9	
Military	20	57.9	
Officers	10	54.3	
Airmen	10	61.5	
Male	28	55.1	
Female	9	41.7	

However, the average error rate and the number of subjects obtaining relatively low error scores show that many subjects were mastering the tasks. If we assume pure guessing among the four choices for each item, we would expect an error score of 75 percent, or 108 errors. No subject scored this poorly. If we assume that only one message was comprehended each trial (that is, no integration of messages), an average of 61 percent, or 88 errors, would be expected (this figure results from an unequal distribution of information between messages). Only two subjects exceeded this error rate, indicating that in general, message pairs were being integrated to obtain answers.

The subgroup means of Table VII suggest a lack of homogeneity in the subject group. The military subgroup contributed the greatest mean error, with airmen accounting for more errors than officers, and this group is also represented in the Hanscom and Male subgroups. There are three possible reasons for this difference between military and student subjects. First, both the Tufts and the Hanscom civilian groups were college students, probably a more homogeneously selected group than the military. Second, the students of both subgroups were volunteers and were research oriented, while the military subjects were assigned to the experiment and were thus, as a group, less likely to be as highly motivated as the students. Third, comments of some subjects suggested that military personnel were troubled by the discrepancies between our experimental exercise and real-life identification procedures and terminology.

Comparing only the male and female student subjects still yielded a mean difference of about 6 errors in favor of female subjects, consistent with the results of Experiment I. However, the number of subjects involved does not warrant the drawing of significant conclusions.

The total group results have been used in the following analysis of the effects of experimental conditions, but the heterogeneity of the subject group must always be kept in mind in drawing conclusions from these results.

<u>Learning Effects</u>: Since all subjects were tested under all conditions (Series A, B and C), increasing familiarity with the task was expected to yield improved performance in successive series. In Table VIII we have computed mean error scores for successive series for all subjects, regardless of the sequence of experimental conditions. Obviously, the expected learning effect occurred, and this, too, must be kept in mind in interpreting the results of the experiment.

<u>Effects of Mode of Presentation</u>: The main effect under study, mode of presentation, was characterized by the Series A, B and C. Series A was the all-visual control series, while Series B and C were the combined audio-visual series. The mean error scores for each condition, regardless of sequence of presentation, are given in Table IX. There is a close similarity in performance between the two bimodal series (B and C), the mean difference of 1.3 errors showing no statistical significance. However, the difference of 3.0 errors between A and C is readily acceptable at the .05 level of significance, whereas the difference of 1.7 between A and B could occur eighteen percent of the time under the null hypothesis, a probability not generally leading to rejection of the hypothesis. Therefore, we thought of the two series B and C as two independent experiments and computed the joint probability of their "t" tests, using the chi-square conversion. The combined chi-square was significant at the 3 percent level of significance, leading to a qualified (because of the heterogeneous group and learning effects) conclusion that performance was somewhat better

TABLE VIII

SUMMARY OF MEAN ERROR SCORES FOR ORDER (N=37), EXPERIMENT II

Order	Mean	S.D.	Difference	"t"	P
First Series	22.2	6.98	1-2	5.69	< 0.001
Second Series	16.2	8.06	1-3	5.97	< 0.001
Third Series	14.0	9.30	2-3	2.71	< 0.05

TABLE IX

SUMMARY OF MEAN ERROR SCORES FOR SERIES (N=37), EXPERIMENT II

Series	Mean	S.D.	Difference	"t"	P
A (Visual)	15.9	9.49	A-B	1.36	0.184
B (Aud-Vis)	17.6	9.03	A-C	2.32	0.026
C (Aud-Vis)	18.9	7.72	B-C	0.74	0.464

under all-visual presentation than under combined audio-visual presentation.

Effects of Sequence of Presentation: The three sequences of presentation of conditions were nearly evenly distributed over the subjects, with the following numbers of cases: Sequence I (ABC) ... 12; Sequence II(BCA)...12; Sequence III(CAB)...13. An examination of errors broken down by sequence (See Table X) is illuminating in view of the heterogeneity of the subject group, for each sequence represents a different subject group. In order to balance for learning effects, we have plotted mean error for each condition in sequential order for each sequence in Figure 7a. The learning effect is evident in the general downward slope with successive series. If we had had comparable subject groups, we would expect an interweaving of these curves to the extent that differences between conditions existed. However, we note instead that Sequence I is consistently higher than the other two. Unfortunately, scheduling problems led to a clustering of the poorest performers (the preponderantly military group) in Sequence I, and the influence of this condition is extremely important in evaluating all results.

Can we estimate how the data would have looked if the group having Sequence I had performed comparably to the other groups? We have subtracted the combined mean error score for Sequences II and III (14.3) from the mean error score for Sequence I (23.8) to obtain a "correction" factor of 9.5, which, applied to Sequence I, yielded three "corrected" points A', B', and C'. These points are plotted in Figure 7b. Then the three points for each condition have been connected, including two actual points and a "corrected" point. This interesting, though speculative, comparison shows a tendency for Series A (all-visual) to yield fewer errors than the bimodal series, with a decreasing difference as

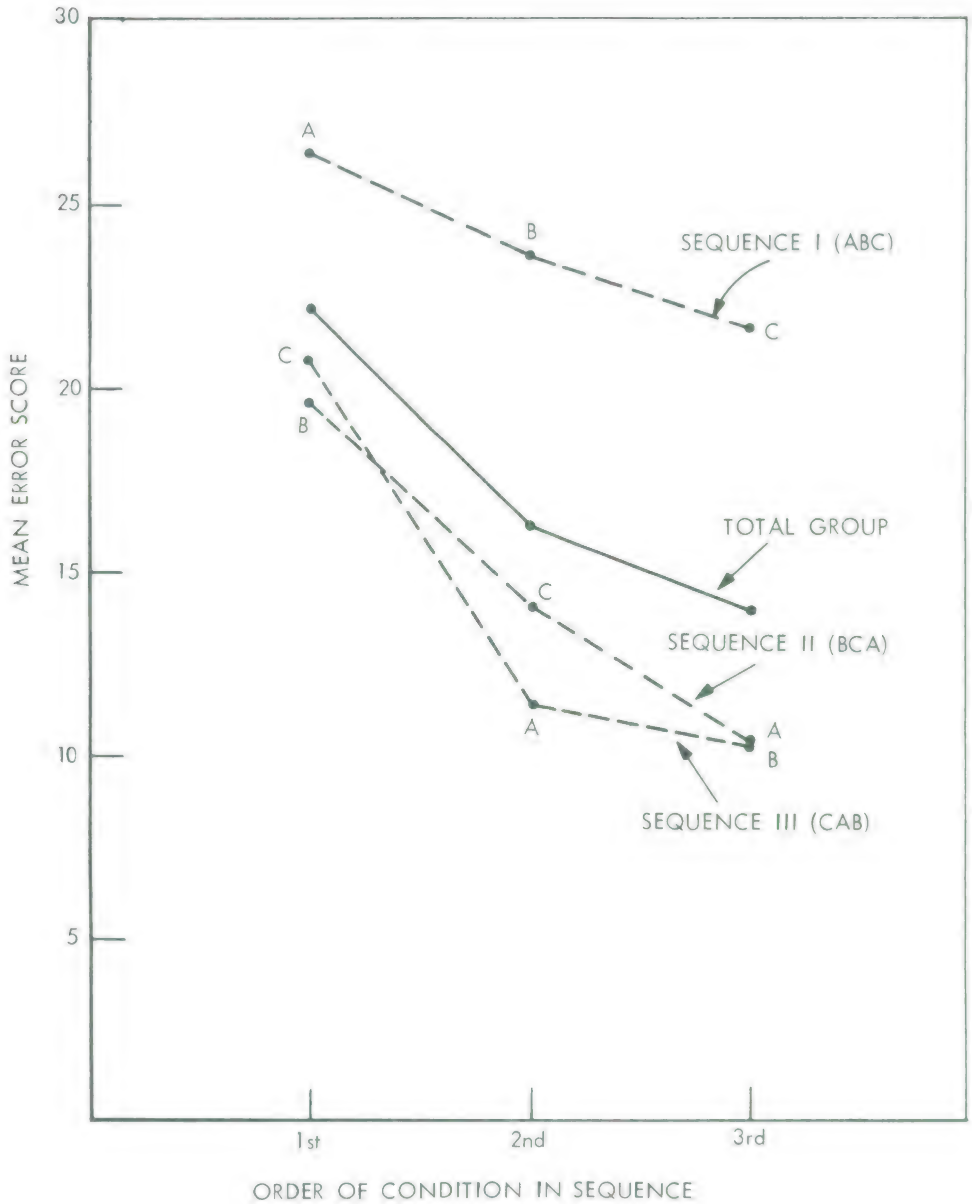

Figure 7a. Mean Errors for Each Series in Each Sequence, Experiment II - Showing Sequence Differences.

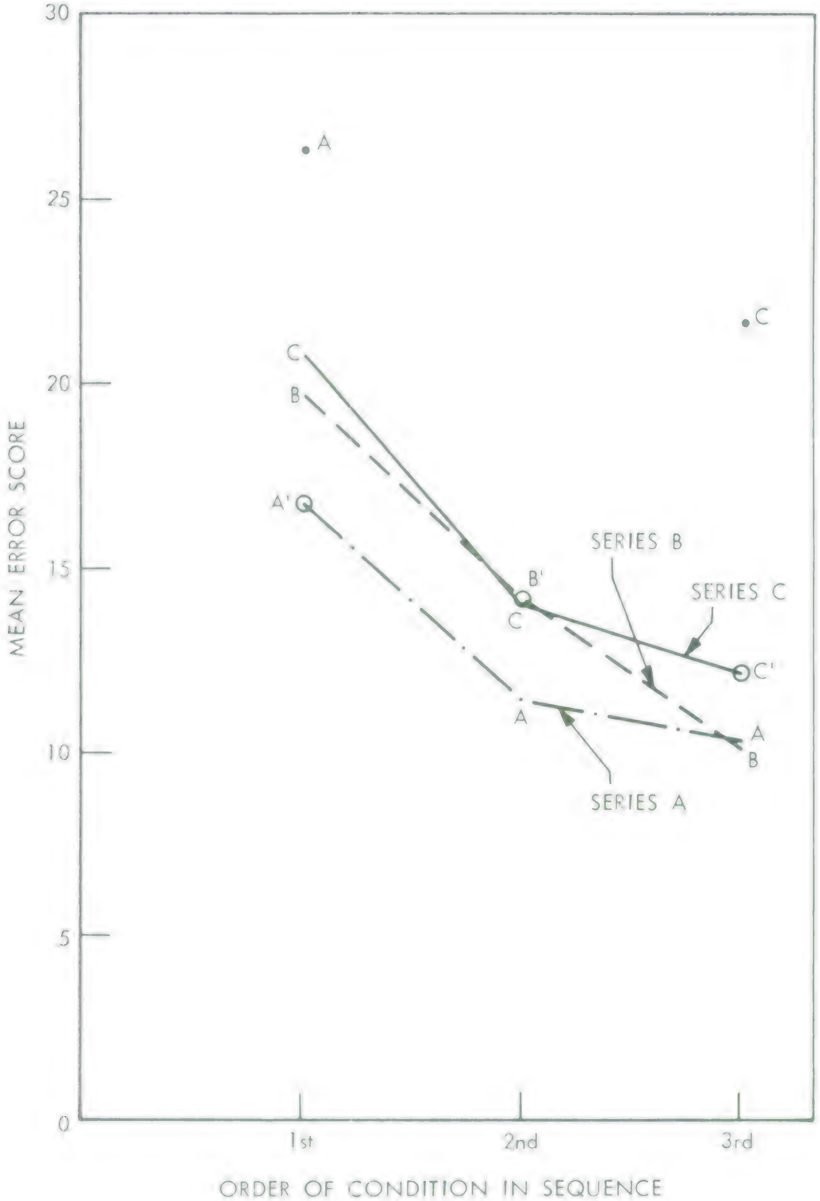

Figure 7b. Mean Errors for Each Series in Each Sequence, Experiment II – Showing Series Results "Corrected" for Sequence I.

TABLE X

MEAN ERRORS FOR EACH SERIES IN EACH SEQUENCE, EXPERIMENT II

Sequence	N	A	B	C	Mean
		Series			
I (ABC)	12	26.3	23.6	21.7	23.8
II (BCA)	12	10.3	19.7	14.0	14.6
III (CAB)	13	11.4	10.2	20.8	14.1
Group	37	15.9	17.6	18.9	

learning continues, strengthening our tentative conclusion made above.

Influence of Practice: One further consideration is necessary. All groups were instructed and given a few practice trials in the all-visual mode. In view of the pronounced improvement with practice between series, we must allow for some influence of this practice on the subsequent series, and this influence would favor Series A. Whether the practice influence is enough to account for the apparent superiority of performance in Series A certainly can not be decided from these data.

Comments of Subjects: The Hanscom subjects were asked to write brief answers to three questions concerning relative difficulty of series, preferred series, and simultaneous versus sequential processing of the two messages. Of the twenty-five subjects responding, fifteen found easier and preferred the all-visual presentation, while ten found easier and preferred the audio-visual presentation. These preferences and subjective evaluations are consistent with the hints in the data that performance was slightly more accurate with the all-visual presentation. Again, however, we can not be sure how much of this feeling in favor of the all-visual presentation resulted from the all-visual examples given in the instructions and pre-test practice.

The subjective judgments on whether the two messages were being processed simultaneously or sequentially divided eighteen-to-seven in favor of sequential processing, a result consistent with the indications of the item analysis of Experiment I ... suggestive but far from conclusive evidence in line with a single-channel theory.

CONCLUSIONS FROM EXPERIMENT II

Experiment II was a difficult task, but performance improved markedly with practice. In the bimodal series, integration of information received simultaneously through different senses certainly did take place. The suggestions in the data that performance was slightly better in the all-visual mode can not be accepted with confidence. On the other hand, there is certainly <u>no</u> evidence of an advantage (in terms of performance) to be gained by using bimodal presentation. Subjective impressions of the subjects were about two-to-one in favor of sequential rather than simultaneous processing of the messages.

SECTION V

CONCLUSIONS AND RECOMMENDATIONS

SUMMARY

Two exploratory experiments were performed to study factors affecting the capability of humans to integrate the information received simultaneously via two different senses. In both experiments, the two senses used as input channels were vision and audition, and verbal messages were used.

In Experiment I, seventy-six college students listened to one message while reading another and then selected the best answer to a question requiring information from both messages. A fifteen-second presentation period was followed by a twenty-second decision period for each of fifty message pairs.

In Experiment II, thirty-seven subjects, both college students and Air Force personnel, listened to one message while reading another in a simulated aircraft identification task and then selected the best identification, the correct answer requiring information from both messages. A two-second presentation period was followed by an eight-second decision period for each of ninety-six bimodal presentations plus forty-eight all-visual control items.

In both experiments, the subjects did, indeed, integrate the bimodally-presented data to select correct answers, with Experiment I proving to be a far easier task than Experiment II. No clear-cut advantage or disadvantage of bimodal versus all-visual presentation was demonstrated.

CONCLUSIONS

The results of these experiments suggest the following primary conclusions:

1. In both experiments, integration of bimodally received information did occur. In Experiment I, the timing was such that attention sometimes could have been alternated between the two senses, but in Experiment II, the stimulation of the two senses had to be concurrent.

2. Comparing experimental results with the results of the all-visual control conditions failed to show any clear-cut advantage or disadvantage to bimodal stimulation. Performance was slightly better (but without statistical significance) for the all-visual condition in Experiment I. A similar result in Experiment II had more statistical weight but could have been accounted for by the initial practice in the all-visual mode.

3. For short messages, at least, learning to integrate bimodally-presented data progresses rapidly. The learning curves of Figure 7 showed just the beginning of leveling and some convergence, suggesting that extensive additional practice might have resulted in still better performance and less intergroup differences. This expectation was partially confirmed by three subjects (not elsewhere reported on) who took part in a preliminary test of Experiment II and were retested during the main running of the experiment. Their performance was generally superior to that of all the other subjects, suggesting considerable retention of skill at the task over a period of three weeks.

With regard to theoretical implications:

4. The results of these experiments have no reliable implications for the single-channel theory of perception. The detailed analysis of three specific items of Experiment I suggested an alternation of attention between senses, which would be consistent with a single-channel system, and the failure of bimodal stimulation to unburden the visual sense would be predicted by a single-channel model. However, the use of two input modes did not measurably increase the difficulty of decision-making, even in Experiment II, where alternation of attention between senses was very unlikely during presentation. A single-channel system <u>could</u> exist, with short-term memory storing one message during processing of the other (supported by a majority of the introspective reports), but nothing in the results suggests that it <u>had</u> to be that way.

The results lead us to some conclusions on experimental methodology:

5. Under the conditions of Experiment I, increasing the number of words or ideas in two messages does not necessarily increase the difficulty of integrating the information in those messages.

6. A mixed group of college students and military personnel can be heterogeneous in the performance of integration of information. Although this is hardly a major discovery, it is well to emphasize the differences not only in selection as a group, but also in motivation and in perception of the experimental situation.

7. Inaccuracies in simulation of a situation can seriously affect the performance of subjects familiar with the real situation. That is, negative transfer in reaction to simulation inaccuracies can outweigh the positive transfer from familiar aspects.

RECOMMENDATIONS

Although this experiment tends to confirm the earlier negative results with regard to the unburdening of one sense by sharing the input load among senses, the question of multimodal presentation of information is far from a dead issue. It seems likely that visual displays and auditory communications will continue to work together in complex control centers. And thus system designers can benefit from principles that guide them in the selection of cues, signals, characters, formats, and the like which will promote integration and prevent interference. The emphasis, however, should be on matching sensory input channels on the basis of compatibility factors rather than load factors.

Our experiments showed that reasonably difficult information processing can occur rapidly and accurately even though parts of the information are received simultaneously through different

senses. On a few dimensions we essentially stacked the cards in favor of getting these results. Several dimensions require exploration now to determine how far along each dimension we can move before integration shifts to interference. Five such dimensions are:

1. Relatedness. The material in our message pairs was always on the same subject matter -- they were meant to go together. How different can we make message contents and still integrate common elements in simultaneous presentation?

2. Appropriateness. We worked with verbal messages, with which people are equally proficient and practiced in vision and audition. Can we expect no deterioration in performance when we assign to a particular sense data for which it is less suited than the "relieved" sense? The failure of Goldman's[28] subjects to track pitch-represented displacements in one dimension along with visually-presented displacements in another dimension may have been partially due to the inappropriateness of audition for representing spatial displacement. Mowbray and Gebhard[5] have catalogued some of the factors for which the various senses are most appropriate, but considerable research must still be done if principles useful in the design of complex control centers are to be derived.

3. Difficulty. What is the relationship between the relative difficulty of two inputs and the ease with which they can be integrated? We know that simple auditory warning signals associate well with complex visual displays. Our attempt in Experiment I to examine difficulty in terms of message length was unsuccessful. In Experiment II we essentially held difficulty between messages constant. Absolute difficulty of the task involved may also help determine whether multimodally presented data will integrate or interfere. More research is required in this area before we can summarize our knowledge in the form of principles.

4. Learning. The rapid improvement with practice shown in Experiment II and the early mastery of the task of Experiment I indicate the importance of practice and learning in situations with multimodal stimulation. Learning should be accounted for in future experimental designs for studies of integration -- it may be a dimension on which interference shifts to integration.

5. Time. Our two experiments tested two presentation times, fifteen seconds and two seconds. Integration occurred in both experiments, although in the longer presentation time we suspect that sometimes alternation of attention resulted in sequential rather than simultaneous stimulation of the two senses. Systematic variation of presentation time should be studied, not so much for itself but for its interactions with message difficulty and length. Related to temporal factors, too, are the fascinating questions of short-term memory, intersensory differences in decay rates, recoding from one sense's information storage to another's and the role of rehearsal. Research on all these questions will undoubtedly show that even the acceptance of a single-channel model of perception does not automatically exclude multimodal stimulation as a technique worthy of study and application in control centers.

REFERENCES

1. Harris, J.D., "Some Relations Between Vision and Audition," U.S. Naval Medical Research Laboratory, New London, Conn. Progress Report No. 1 on BuMed Project NM000-009, July, 1948.

2. Cheatham, Paul G., "A Comparison of the Visual and Auditory Senses as Possible Channels for Communication," Air Materiel Command, Wright-Patterson Air Force Base, Dayton, Ohio, A.F. Tech. Report No. 5919, May, 1950.

3. Day, W.F. and Beach, Barbara R., "A Survey of the Research Literature Comparing the Visual and Auditory Presentation of Information," Air Materiel Command, Wright-Patterson Air Force Base, Dayton, Ohio, A.F. Tech. Report No. 5921, November, 1950.

4. Henneman, R.H. and Long, E.R., "A Comparison of the Visual and Auditory Senses as Channels for Data Presentation," Aero Medical Laboratory, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, WADC TR 54-363, August, 1954.

5. Mowbray,G.H. and Gebhard, J.W., "Man's Senses as Information Channels," Johns Hopkins University, Applied Physics Laboratory, Silver Spring, Maryland, Report No. CM-936, May, 1958.

6. Sinaiko, W.H. (Ed.) Selected Papers on Human Factors in the Design and Use of Control Systems, New York: Dover, 1961.

7. London, I.D., "Research on Sensory Interaction in the Soviet Union," Psychol. Bull., 1964, 51, 531-568.

8. San Diego State College Foundation, "Bibliography - Comparison and Interaction Among Sensory Input Channels," Office of Naval Research, Contract NONR-1268(01), March 1, 1955.

9. Broadbent, D.E., "Failures of Attention in Selective Listening," Medical Research Council, Cambridge, England, A.P.U. 168/51, 1951.

10. Broadbent, D.E., "Listening Between and During Practiced Auditory Distractions," Medical Research Council, Cambridge, England, A.P.U. 208/54, 1954.

11. Broadbent, D.E., "A Review of Multiple Channel Listening Experiments," Medical Research Council, Royal Navy Personnel Research Committee, England, RNP 55/833, April, 1955.

12. Broadbent, D.E., "Successive Responses to Simultaneous Stimuli," Quart. J. Exp. Psychol., 1956, 8, 145-152.

13. Buckner, D.N. and McGrath, J. J., "A Comparison of Performances on Single and Dual Sensory Mode Vigilance Tasks," Personnel and Training Branch, Office of Naval Research, Department of the Navy, NR 153-199, February, 1961.

14. Gruber, A., "Sensory Alternation and Performance in a Vigilance Task," Decision Sciences Laboratory, L.G. Hanscom Field, Bedford, Mass., ESD-TDR-63-605, September, 1963.

15. Loveless, N.E., "Signal Detection with Simultaneous Visual and Auditory Presentation," Flying Personnel Research Committee, London, England, FPRC 1027, December, 1957.

16. Schafer, T.H. and Shewmaker, C.A., "A Comparative Study of the Audio, Visual and Audio-Visual Recognition Differentials for Pulses Masked by Random Noise," U.S. Naval Electronics Laboratory, San Diego, California, Tech. Report No. 372, 1953.

17. Stirner, F.W., Siegel, A.I. and Baker, R.C., "Caution and Warning Light Indicators for Naval Aircraft: V. an Experimental Comparison of Visual, Auditory and Visual-Auditory 'Master' Signals," USN Aircrew Equipment Lab., Air Materiel Center, Philadelphia, Penn., NAMC ACEL 357, TED NAM EL 52004, Part 17, December, 1957.

18. Mudd, S.A. and McCormick, E.J., "The Use of Auditory Cues in a Visual Search Task," J. Applied Psychol., 1960, 44, 184-188.

19. Sumby, W.H. and Pollack, I., "Visual Contribution to Speech Intelligibility in Noise," J. Acoust. Soc. Am., 1954, 26, 212-215.

20. Decker, L.R. and Pickett, J.M., "Visual Indication of Speech as an Aid to Hearing It in Noise and Babble," J. Acoust. Soc. Am., 1957, 29, 150-151.

21. Tolhurst, G.C., "The Effects of Disrupting the Simultaneity of Visual-Aural Communication Channels to a Speaker," USN School of Aviation Medicine, Pensacola, Florida, NM 18 02 99 REP 66, December, 1956.

22. Zuckerman, J.V., "Commentary Variations: Level of Verbalization, Personal Reference, and Phase Relations in Instructional Films on Perceptual-Motor Tasks," Special Devices Center, Office of Naval Research, Tech. Report-SDC-269-7-4, December, 1949.

23. Nelson, H.E. and Moll, K.R., "Comparison of the Audio and Video Elements of Instructional Films," J. Communication, 1951, 1, 62-66.

24. Nelson, H.E. and Vander Meer, A.W., "The Relative Effectiveness of Differing Commentaries in an Animated Film on Elementary Meteorology," Special Devices Center, Office of Naval Research, Tech. Report, SPECDEVCEN 269-7-43, June, 1955.

25. Elliot, F.R., "Memory for Visual, Auditory and Visual-Auditory Material," Archives of Psychology, 1936, 29 (199), 5-57.

26. Forbes, T.W., Garner, W.R. and Howard, J.G., "Flying by Auditory Reference," National Defense Research Commission, Washington, D.C., OSRD Report No. 5123, 1945.

27. Mowbray, G.H., "The Perception of Short Phrases Presented Simultaneously for Visual and Auditory Reception," Quart. J. Exp. Psychol., 1954, 6, 86-92.

28. Goldman, A., "Sensory Interaction and Response Capacity," U.S. Naval Training Device Center, Port Washington, New York, Tech. Report: NAVTRADEVCEN 789-11-1, June, 1959.

29. Fitts, P.M., "Engineering Psychology and Equipment Design," In S.S. Stevens (Ed.) Handbook of Experimental Psychology, New York: Wiley, 1951, Ch. 35, 1287-1340.

30. Davis, R., "The Human Operator as a Single Channel Information System," Quart. J. Exp. Psychol., 1957, 9, 119-129.

31. Kristofferson, A.B., "The Psychological Time Quantum and the Discrimination of Succession," Bolt, Beranek and Newman, Inc., Cambridge, Mass., Contract NAS2-2486, Report No. 1233, March, 1965.

32. Adams, J.A. and Creamer, L.R., "Data Processing Capabilities of the Human Operator," J. Eng. Psychol., 1962, 1, 149-158.

33. Sperling, G., "A Model for Visual Tasks," Human Factors, 1963, 5, 19-31.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Applied Research Lab. Sylvania Electronic Systems, A Division of Sylvania Electric Products Inc., 40 Sylvan Road, Waltham, Massachusetts 02154	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP N/A

3. REPORT TITLE

Integration of Concurrent Visual and Auditory Messages

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Final Report

5. AUTHOR(S) *(Last name, first name, initial)*

Devoe, Donald B.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 1965	62	33

8a. CONTRACT OR GRANT NO. AF19(628)-4073	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. BPSN 6899-7802-7682	ESD-TR-65-461
c. TASK 768201	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	F-4068-1

10. AVAILABILITY/LIMITATION NOTICES Qualified requestors may obtain copies of this report from DDC. Other persons or organizations should apply to the Clearinghouse for Federal Scientific and Technical Information (CFSTI), Sills Building, 5285 Port Royal Rd., Springfield, Va. 22151

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Decision Sciences Lab. (ESRHD) Electronic Systems Division Air Force System Command, USAF L. G. Hansom Field, Bedford, Mass.

13. ABSTRACT

Two experiments were performed involving the concurrent presentation to human subjects of two messages, one auditory and one visual, followed by a question requiring information from both messages. The results indicated that bimodally-presented information can be integrated for decision making. However, there was no evidence of an advantage to bimodal presentation as a means of unburdening an overloaded sense. The implications of the results for displays and communications in complex control centers are discussed and directions for future research are suggested.

DD FORM 1473 1 JAN 64

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
AUDIO-VISUAL						
AUDITION						
DISPLAY						
INTEGRATION						
INTERFERENCE						
MESSAGE						
PERCEPTION						
UNBURDENING						
VISION						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, *or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.